

# Реализация глобальной климатической модели на многопроцессорной ЭВМ кластерного типа\*

В.П. Пархоменко

Реализованы варианты параллельного алгоритма климатической модели для различных способов разбиения расчетной области с модифицированными коммуникационными процедурами для обменов информацией между процессорами в параллельном варианте модели общей циркуляции атмосферы (ОЦА). Представлен анализ некоторых факторов, влияющих на выполнение параллельной реализации модели ОЦА на многопроцессорной ЭВМ кластерного типа. Обсуждаются несколько модификаций первоначального параллельного кода модели ОЦА, направленных на улучшение его вычислительной эффективности, баланса загрузки процессоров.

## 1. Введение

Климатическая система Земли характеризуется сложными взаимодействиями и обратными связями ее элементов: атмосфера, океан (с морским льдом) и деятельный слой суши (почва и растительность). Модели общей циркуляции атмосферы (ОЦА) и океана - среди наиболее мощных инструментов, используемых для изучения климатической системы и ее изменчивости. Численное моделирование, выполненное с использованием таких моделей, требует значительных вычислительных затрат, потому что большое количество трехмерных физических полей должно быть определено на каждом временном шаге при решении системы нелинейных уравнений в частных производных, описывающих движение жидкости на вращающейся сфере, и также потому что требуется очень длинный период моделирования, чтобы получить статистически значимые числовые результаты. Параллельные компьютеры, таким образом, естественные инструменты для моделирования климата [1].

Параллельное выполнение модели показало [2], что эффективность параллельного кода при большом числе процессоров ухудшается. Другими словами, программа плохо "масштабируется" с увеличением числа процессоров.

Одна из целей работы состоит в том, чтобы проанализировать компоненты алгоритма модели ОЦА и их параллельное выполнение, найти узкие места, которые препятствуют эффективности параллельного алгоритма, использовать лучшие алгоритмы и более эффективные параллельные стратегии выполнения, увеличить скорость выполнения модели ОЦА на параллельных системах [2].

Предлагаемая модель включает атмосферный блок, реализованный на базе модели общей циркуляции атмосферы с параметризацией ряда подсеточных процессов. Разработаны версии модели с более мелким пространственным разрешением и моделью общей циркуляции океана [3]. Взаимодействие между блоками осуществляется в интерактивном режиме. Система трехмерных дифференциальных уравнений модели атмосферы, так называемая система примитивных уравнений [4], включает в себя уравнение, описывающее гидростатическое приближение по вертикали, прогностическое уравнение, описывающее закон сохранения массы сухого воздуха, прогностическое уравнение для сохранения влагосодержания в атмосфере, динамические прогностические уравнения для горизонтальных компонент скорости и термодинамическое уравнение сохранения энергии. Уравнения дополняются соответствующими граничными условиями и, таким образом, получается замкнутая динамическая система.

По вертикали атмосфера разбита на девять уровней, от поверхности до высоты, соответствующей давлению 200 мбар. В качестве вертикальной принята сигма система координат, связанная с гидростатическим давлением. Для определения источников водяного пара и тепла применяются модели, описывающие гидрологический цикл и процессы

---

\* Работа выполнена при поддержке Программы фундаментальных исследований Президиума РАН № 14 и Проекта РФФИ №08-01-00607.

распространения теплового и солнечного излучения. Явная численная схема, используемая в модели, накладывает ограничение сверху на шаг интегрирования по времени, который даже при данном пространственном разрешении модели не превышает 10 мин.

Испарение, конденсация и конвекция зависят от термического состояния атмосферы, которое в свою очередь является функцией обмена теплом, имеющим место в этих процессах. Вместо получения одновременного решения задач для определения влаги и термического состояния, в модели рассчитываются испарение и компоненты конденсации последовательно. На каждом шаге модифицируется термическое состояние атмосферы, новые значения температуры используются на следующем шаге. При этом применяется так называемая процедура конвективного приспособления [4].

Модель имеет относительно грубое пространственное разрешение, однако при сравнительно невысоких вычислительных затратах позволяет исследовать механизмы и способы распараллеливания для исследования их эффективности и определения узких мест. Разрешающая способность глобальной модели может быть увеличена. Она зависит от мощности используемых ЭВМ.

## 2. Структура и выполнение параллельного кода модели ОЦА

Модель ОЦА - пакет программ, который моделирует много физических процессов [3]. Можно выделить два главных компонента программы: динамический блок, который вычисляет течения жидкости, описываемые примитивными уравнениями посредством конечных разностей, и физический блок, который вычисляет влияние мелкомасштабных процессов (типа конвекции в масштабах облака) на процессы, которые разрешаются сеткой. Результаты, полученные в физическом блоке, используются в динамическом блоке как внешние источники для вычисления течения. Программа модели ОЦА использует трехмерную разнесенную сетку для скоростных и термодинамических переменных (потенциальная температура, давление, удельная влажность, и т.д.). Эта трехмерная сетка является C-сеткой Аракавы [3] в горизонтальном (широта/долгота) направлении с относительно небольшим числом вертикальных слоев (обычно значительно меньше, чем число горизонтальных точек сетки). Динамический блок состоит из двух главных компонент: фильтрации и собственно конечно-разностных вычислений. Операция фильтрации необходима на каждом временном шаге в областях близких к полюсам, чтобы гарантировать, что эффективный размер сетки там удовлетворяет условию устойчивости для явной схемы по времени [4].

В параллельной реализации модели используется разбиение по процессорам двумерной сетки в горизонтальной плоскости. Этот выбор основан на том, что вертикальные процессы сильно связывают точки сетки, что делает распараллеливание менее эффективным в вертикальном направлении, и на том, что число точек сетки в вертикальном направлении является обычно небольшим. Каждая подобласть в такой сетке - прямоугольная область, которая содержит все точки сетки в вертикальном направлении. В этом случае есть в основном два типа межпроцессорных обменов [2]. Обмены данными необходимы между логически соседними процессорами (узлами) при вычислениях конечных разностей (задействованы только соседние точки сетки); удаленные обмены данными необходимы, в частности, чтобы осуществить операции спектрального фильтрации. Соотношение затрат времени главных блоков исходной параллельной программы модели ОЦА, при использовании  $4^{\circ} \times 5^{\circ} \times 9$  разрешения, которое содержит  $46 \times 72 \times 9$  точек, показано в Таблице 1.

**Таблица 1.** Соотношение затрат времени счета основных блоков модели

Число процессоров	Динамический блок, %	Физический блок, %
1	63	33
8	67	30
16	70	27

Как следует из Таблицы 1, основная часть вычислительных затрат модели ОЦА связана с блоком динамики и блоком физики, с исключенными процедурами ввода - вывода. Эти процедуры выполняются только один раз, тогда как главная часть повторяется многократно по

времени и доминирует по затратам времени выполнения. Сравнивая эти два блока, мы можем видеть, что динамическая часть занимает основное время счета, особенно при большом количестве узлов. На масштабируемость параллельной программы влияет отношение затрат обменов данными к затратам вычислений и степень несбалансированности загрузки процессоров в программе. Результаты анализа и оценок затрат времени показывают, что затраты спектральной процедуры - заметная часть в параллельной программе модели ОЦА, особенно при увеличении числа узлов.

### 3. Равномерная загрузка процессоров в процедуре сглаживания

Два типа фильтрации выполняются в программе ОЦА. Один – трехточечное сглаживание, которое применяется к широтам от  $38^\circ$  до  $70^\circ$  в каждом полушарии; другой - набор дискретных Фурье фильтров, специально предназначенных, чтобы демпфировать быстрые инерционно-гравитационные волны около полюсов. Фильтры содержат широтную зависимость, но применяются во всей долготной области на каждом вертикальном слое от полюсов до  $70^\circ$  в каждом полушарии. В остальной области фильтрация не применяется.

Поскольку в преобразовании Фурье задействованы все точки широтного пояса, то используемое в программе распараллеливание не будет эффективным. Поэтому для преобразования Фурье используется метод распараллеливания, в котором все частичные суммы вычисляются на своем процессоре, а затем уже на всех процессорах собираются полные суммы. Данный алгоритм используется как при вычислении нормировочных коэффициентов (которое производится один раз при первом вызове процедуры сглаживания), так и непосредственно в преобразовании Фурье. Это позволяет эффективно распараллелить данный участок программы.

Высокие затраты фильтрации, по сравнению с остальной частью динамического модуля (Таблица 2), являются также следствием того, что существует заметная неравномерность загрузки процессоров, так как только подобласти в высоких широтах требуют фильтрации. Для решения этих проблем необходимо выполнять балансировку загрузки перед фильтрацией, перераспределяя данные, подвергающиеся сглаживанию с процессоров, содержащих подобласти высоких широт на процессоры, содержащие подобласти низких широт, которые или выполняют трехточечное сглаживание, или полностью простаивают в течение стадии фильтрации.

**Таблица 2.** Относительные временные затраты фильтрации Фурье в динамическом блоке

<b>Число процессоров</b>	1	8	16
<b>Фильтрация Фурье, %</b>	10	18	24

Чтобы решить проблему равномерной загрузки при фильтрации, нужно перераспределить ряды данных, которые будут фильтрованы по широтному направлению. В программе ОЦА спектральное фильтрация выполняется на каждом временном шаге перед вызовом конечно-разностных процедур. Фильтрация выполняется для различных наборов физических переменных, последовательно по одной переменной в исходной программе ОЦА. Чтобы максимизировать эффективность выполнения программы в зависимости от равномерности процедуры загрузки, предлагается реорганизовать процесс фильтрации так, чтобы все переменные были фильтрованы одновременно. Это изменение возможно, потому что данные независимы. Основываясь на этих соображениях, логично создать общий модуль балансировки загрузки, который делает следующее: данную  $M \times N$  сетку процессоров, с  $M$  процессорами в широтном направлении и  $N$  процессорами в долготном направлении, с  $L$  переменными, которые будут фильтрованы, каждый с  $R_j$  ( $j = 1, \dots, L$ ) рядами данных, которые будут фильтрованы, перераспределяет ряды данных в долготном направлении так, чтобы после перераспределения, каждый процессор содержал приблизительно (так как общее количество

рядов данных, которые будут фильтрованы обычно не делится на  $N$ )  $\left( \sum_{j=1}^L R_j \right) / N$  рядов,

которые будут фильтрованы. Если предположить, что точно половина рядов данных в одном полушарии должна быть фильтрована, выполнение перераспределения данных для равномерной загрузки было бы относительно простой задачей. Все, что требовалось бы в этом случае, это перераспределить ряды данных способом, который является симметричным относительно линии широты  $45^\circ$  в каждом полушарии. Перераспределение загрузки сопровождается перемещением ряда данных, и перераспределением среди процессоров в долготном направлении. Фильтрация Фурье выполняется на рядах данных после перемещения данных, которое затем сопровождается обратными перемещениями данных, чтобы восстановить расположение данных, которое существовало до фильтрования. Некоторая программа установки необходима, чтобы получить информацию, которая управляет перемещением данных для балансирования загрузки и сбалансированного по загрузке параллельного преобразования Фурье. Программа установки включает фактические подсчеты и управляет межпроцессорными коммуникациями. Соответствующие временные затраты незначительны при моделировании ОЦА, так как это делается только один раз, и они почти не зависят от размера задачи ОЦА.

#### 4. Модификация первоначального параллельного кода модели для улучшения его вычислительной эффективности.

Как было отмечено, в модели ОЦА можно выделить два главных компонента: динамический блок и физический блок (блок источников). Результаты, полученные в физическом блоке, используются в динамическом блоке как внешние источники для вычисления течения.

Подлежащие интегрированию по времени прогностические уравнения для основных искомым функций (горизонтальных компонент скорости, температуры, отношения смеси водяного пара) можно записать в точке горизонтальной конечно-разностной сетки в виде

$$\frac{\partial \psi}{\partial t} = D(\psi) + S(\psi),$$

где  $\psi$  – любая из основных искомым функций.

Здесь через  $S(\psi)$  обозначены «источники» (определяемые в блоке физики), к которым отнесены сила трения в уравнении движения, адиабатическое нагревание в уравнении энергии и источники влаги в уравнении переноса водяного пара. Эти члены являются локальными, то есть не содержат производных по горизонтали. Все остальные («динамические») члены включены в  $D(\psi)$  и содержат производные по горизонтали.

Для решения прогностических уравнений в исходном скалярном варианте программы используется метод расщепления по физическим процессам. Он представляет собой явный шестишаговый метод интегрирования по времени, при котором сначала выполняется шесть шагов по времени с шагом  $\Delta t_D$  с учётом влияния только динамических членов с различной аппроксимацией пространственных производных, а затем один шаг  $\Delta t = \beta t_D$  с учётом влияния только источников (блок физики).

Таким образом, концептуально исходная схема интегрирования по времени такова:

$$\text{I этап: } \psi_D^1 = \psi^0 + \Delta t D(\psi^0),$$

$$\text{II этап: } \psi^1 = \psi_S^1 = \psi^0 + \Delta t S(\psi_D^1),$$

где  $\psi^0$  – значение  $\psi$  в момент времени  $t$ ,  $\psi_D^1$  – в момент времени  $t + \Delta t$  после учёта динамических членов,  $\psi^1 = \psi_S^1$  – окончательное значение  $\psi$  (после учёта влияния источников) в момент времени  $t + \Delta t$ . Совокупность I и II этапов обеспечивает аппроксимацию рассматриваемых уравнений.

Предлагаемый метод распараллеливания предполагает одновременный расчёт вклада физики и динамики соответственно на двух группах процессоров. Реализация метода требует изменения численной схемы по времени. Оно состоит в одновременном расчете динамики и физики на различных группах процессорах с одних и тех же входных данных.

Таким образом, модифицированная схема расчета такова:

$$\text{I группа процессоров: } \psi_D = \psi^0 + \Delta t D(\psi^0),$$

$$\text{II группа процессоров: } \psi_S = \psi^0 + \Delta t S(\psi^0).$$

По окончании циклов мы получаем значения основных переменных  $\psi_D$  на первой группе процессоров на следующем шаге по времени и  $\psi_S$  на второй группе процессоров на следующем шаге по времени. После этого процессоры обмениваются данными и на каждой из групп процессоров по формуле

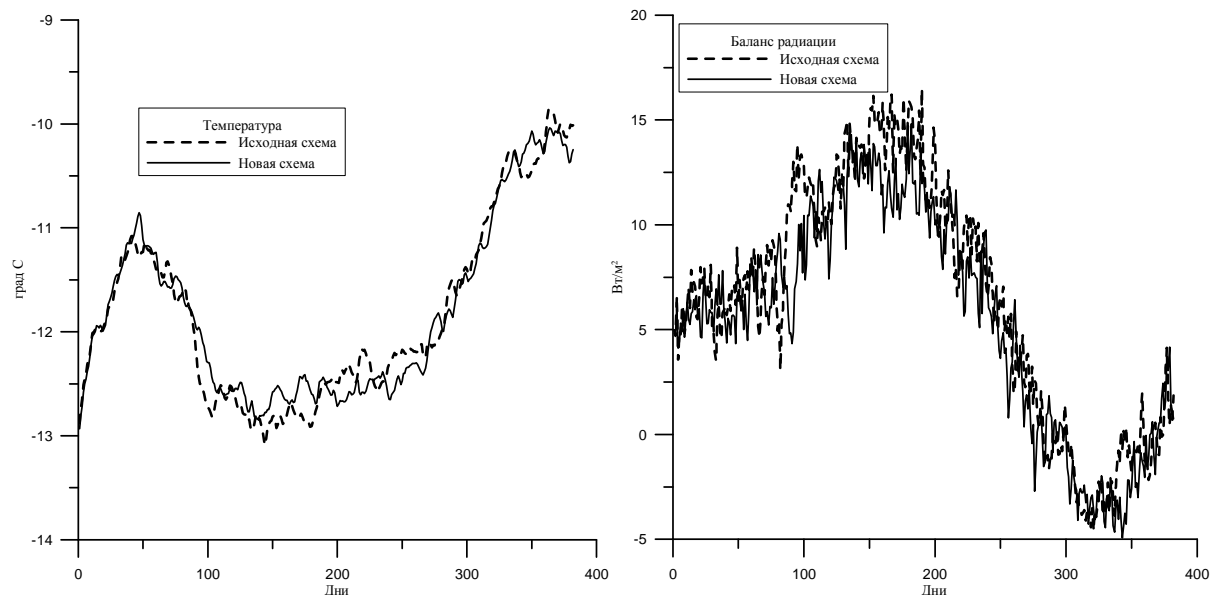
$$\psi^1 = \psi_D + \psi_S - \psi^0$$

рассчитываются окончательные значения искомых функций в момент времени  $t + \Delta t$ . Легко видеть, что при этом достигается аппроксимация уравнений, но понятно, что результаты исходной и модифицированной схем будут отличаться.

В качестве средства реализации распараллеливания использована библиотека MPI. В настоящее время MPI является наиболее распространенным средством распараллеливания, его реализации есть практически на всех современных многопроцессорных вычислительных машинах, что позволяет обеспечить переносимость программы [1].

Исходная программа была модифицирована в соответствии с высказанными выше соображениями. Для проверки корректности метода: по модифицированному и старому вариантам программы были проведены модельные расчеты до установления. Они начинались с одинаковых начальных условий.

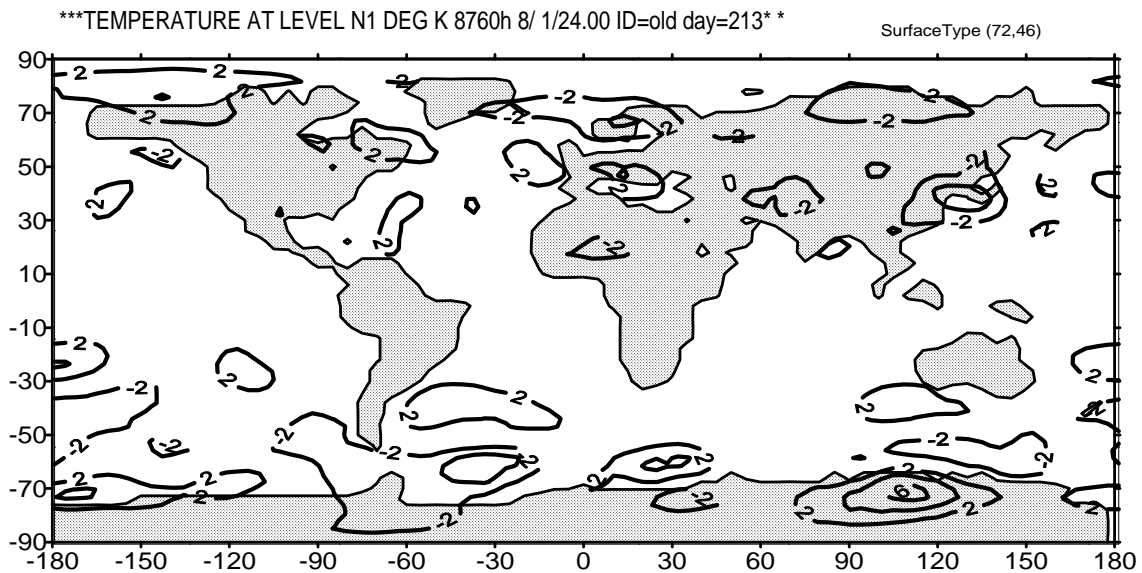
Далее на рисунках приведены некоторые результаты сравнения. На рис. 1 показаны зависимости средней температуры и баланса радиации на верхней границе атмосферы, от времени для исходной и модифицированной схем. Эти средние за сутки и по всему земному шару величины демонстрируют хорошее совпадение результатов, учитывая сильную изменчивость характеристик атмосферы на коротком интервале в одни сутки.



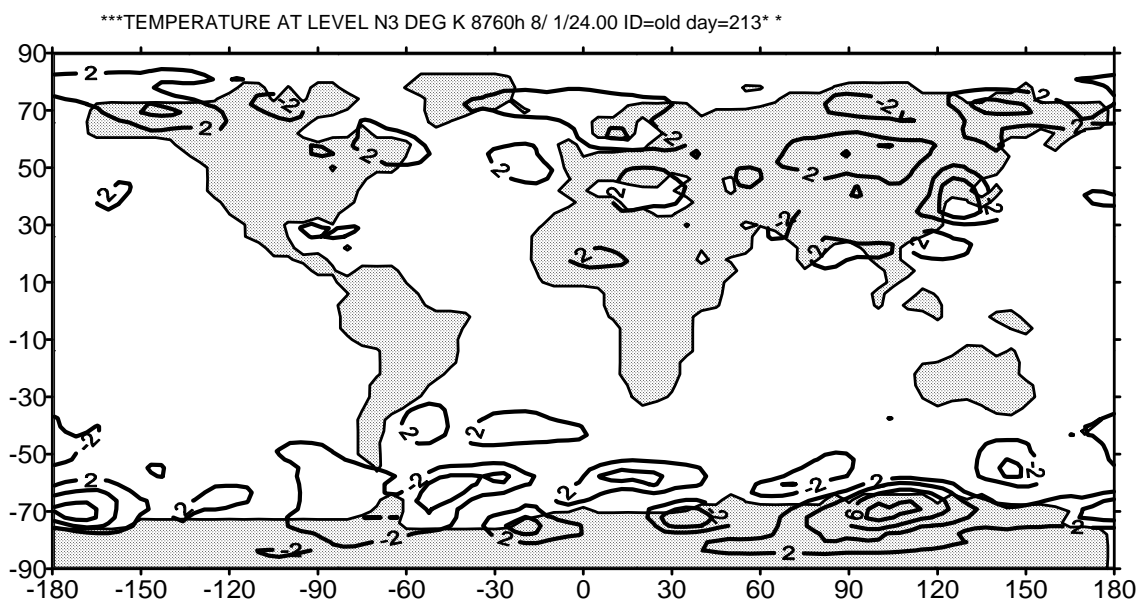
**Рис. 1.** Зависимости средних глобальных температуры (слева) и баланса радиации (справа) на верхней границе атмосферы от времени для исходной и модифицированной схем

Анализ результатов демонстрирует очень хорошее совпадение для обеих схем зональной компоненты скорости и приземной температуры воздуха, осредненных по долготе для июля и января месяцев. Скорости имеют отличия в зимний период в областях сильных градиентов.

На рис. 2, 3 показано распределение разности температур воздуха для июля месяца на уровнях 400 мбар и 800 мбар соответственно. На уровне 400 мбар отличия не превышают всюду 2°C, за исключением одной точки, где составляют 6°C. На уровне 800 мбар картина схожая. Можно утверждать, что отличия в результатах более существенны в приземных областях.



**Рис. 2.** Изолинии разностей температуры воздуха на уровне 400 мбар, рассчитанной по исходной и модифицированной схемам. Июль месяц.



**Рис. 3.** Изолинии разностей температуры воздуха на уровне 800 мбар, рассчитанной по исходной и модифицированной схемам. Июль месяц.

Отметим общую особенность распределений основных климатических характеристик, рассчитанных по исходной и модифицированной схемам – наибольшие отличия результатов наблюдаются в средних и высоких широтах в зимний период. Это связано, по-видимому, с выпадением осадков в виде снега, и интенсивных нестационарных конвективных процессов в

атмосфере. Отличия в приземных температурах в более чем 90% ячеек составляют менее 2°C. Только в двух ячейках зимой в Северном полушарии над материком разница составляет 10°C. Для зимы в Южном полушарии наблюдается похожая картина: заметные отличия существуют в трех ячейках в Антарктиде.

Отличия в приземном давлении не превышают 15 мбар в нескольких ячейках и в основном составляют менее 5 мбар.

Отметим также, что отличия в результатах двух схем не являются препятствием для использования любой из них для прогнозирования изменения климата, так как при расчетах такого рода проводится сравнение прогнозируемого климата и современного для одной и той же схемы. Характер и количественные показатели изменения модельного климата в основном определяются описанием климатических процессов в модели, а они одинаковы в обеих схемах. В настоящее время все климатические модели в той или иной степени дают ошибки при описании локальных характеристик и разницу в прогнозах.

Таким образом, анализ показывает, что результаты расчетов по модифицированной расчетной схеме дают удовлетворительные результаты и возможно ее применение. В нераспараллеленной программе время счета блока физики составляет 38 %, а время счета блока динамики – 62 %. Это означает, что на распараллеленной программе может быть достигнуто ускорение в полтора раза (при использовании двух процессоров). Предлагаемая процедура используется вместе с процедурами распараллеливания блоков динамики и физики на основе декомпозиции расчетной области. Здесь можно применить разработки, описанные в [2, 5, 6]. Это позволяет оптимизировать загрузку процессоров и повысить эффективность распараллеливания. На рис. 4 показаны наилучшие результаты по ускорению при применении описанных подходов. Для счета был использован кластер МВС-6000IM (256 CPU, 64-разрядные процессоры Intel® Itanium-2® 1.6 ГГц). Вычислительные модули связаны между собой высокоскоростной коммуникационной сетью Myrinet (2 Гбита/сек).

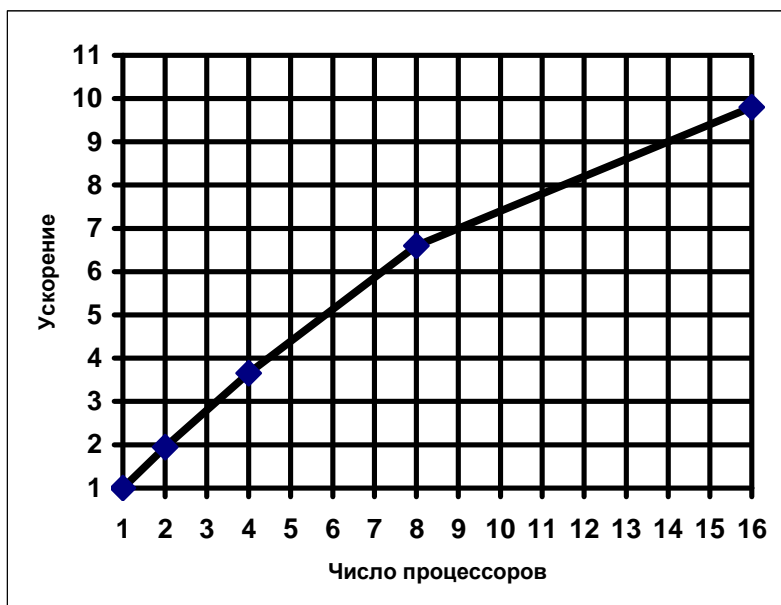


Рис. 4. Зависимость ускорения от числа процессоров.

Другая возможность применения метода – усложнение блока физики, без увеличения общего времени счета. Проведены также долговременные расчеты с целью исследования естественной изменчивости арктического морского льда.

## 5. Моделирование естественной изменчивости арктического морского льда

В термодинамической модели эволюции морского льда [3], как блока климатической модели лед рассматривается как ровная однородная пластина, лежащая на воде и покрытая однородным слоем снега. Учитываются процессы нарастания и таяния льда, а также образования внутренних проталин в толще льда.

Расчеты базового варианта для современного климата начинаются с начальных условий ледового покрова Северного Ледовитого океана (СЛО), когда его толщина всюду постоянна и равна 2 м, сплоченность равна 0.9 и снеговой покров отсутствует. Площадь и конфигурация ледового покрова соответствует климатическим данным наблюдений для соответствующего месяца. В процессе выхода на установившийся режим сначала происходит быстрое летнее таяние льда в окраинных областях, что приводит к уменьшению средней толщины льда. Намерзание и накопление льда в центральных областях Арктического бассейна в зимний период происходит медленно, поэтому выход на стационарный режим происходит постепенно и занимает около 15 лет. На рис. 5 показано изменение во времени средней толщины льда в СЛО. Жирной линией обозначено скользящее среднее этой величины с периодом осреднения 1 год. В установившемся режиме присутствуют значительные межгодовые колебания ледового покрова.

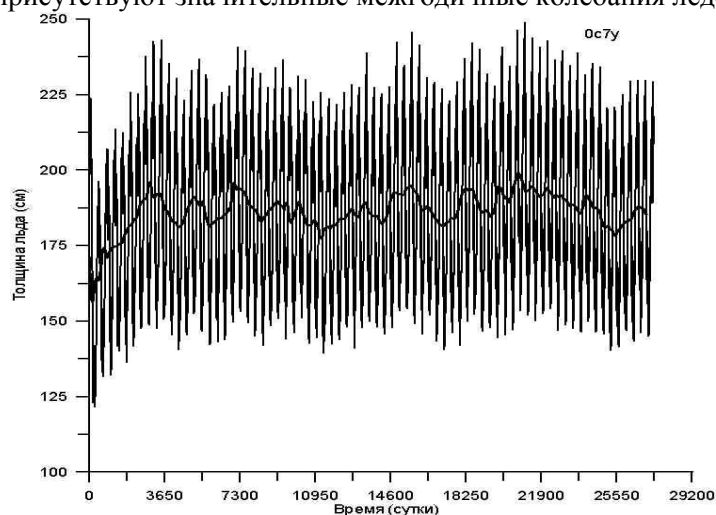


Рис. 5. Изменения во времени средней толщины льда в СЛО. Результаты расчетов

На рис. 6 представлены распределения толщины ледового покрова в Арктике для марта (максимальная площадь ледового покрова) и сентября (минимальная площадь ледового покрова) месяцев.

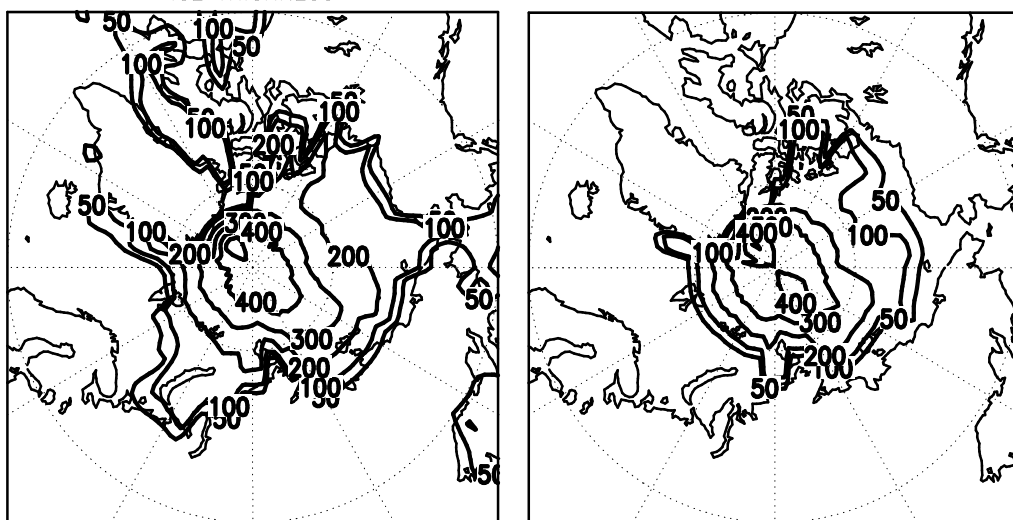


Рис. 6. Толщина льда (см) в СЛО в марте и сентябре. Результаты расчетов



Численные эксперименты позволили установить сложный характер зависимости ледового покрова от параметров модели. Так, уменьшение альбедо снега и льда ведет к уменьшению средней толщины льда, как и следовало ожидать. Однако существуют области, где происходит рост толщины ледового покрова. Аналогичная ситуация и с другими параметрами.

## 6. Заключение

Реализованы варианты параллельного алгоритма климатической модели для различных способов разбиения расчетной области с модифицированными коммуникационными процедурами для обменов информацией между процессорами в параллельном варианте модели общей циркуляции атмосферы. Осуществлена реализация параллельной программы для различных способов разбиения расчетной области по процессорам в климатической модели. Проведены тестовые расчеты для сравнения с однопроцессорным вариантом

Представлен анализ некоторых факторов, влияющих на выполнение параллельной реализации модели общей циркуляции атмосферы на многопроцессорной ЭВМ кластерного типа. Обсуждаются несколько модификаций первоначального параллельного кода модели ОЦА, направленных на улучшение его вычислительной эффективности, баланса загрузки процессоров. Обнаружено, что выполнение сбалансированной загрузки алгоритма сглаживания Фурье обеспечивает сокращение времени выполнения приблизительно на 40%, по сравнению с первоначальным алгоритмом на основе свертки.

Проведена модификация численной схемы по времени модели общей циркуляции атмосферы для возможности осуществления параллельных расчетов блоков динамики и физики с оценкой эффективности распараллеливания. Предлагаемая процедура используется вместе с процедурами распараллеливания блоков динамики и физики на основе декомпозиции расчетной области. Это позволяет оптимизировать загрузку процессоров и повысить эффективность распараллеливания. Другая возможность применения метода – усложнить блок физики, без увеличения общего времени счета.

Были проведены тестовые расчеты на высокопроизводительном кластере. Получено удовлетворительное совпадение результатов для однопроцессорной и параллельной программ: заметные отличия появляются только через сотни шагов по времени, что связано с процедурами округления. Проведены долговременные расчеты с целью исследования естественной изменчивости арктического морского льда.

## Литература

1. Воеводин В.В., Воеводин Вл.В. Параллельные вычисления. СПб.: БХВ-Петербург, 2002. 600 с.
2. Забелок С.А., Пархоменко В.П., Шипилин А.В. Разработка и реализация параллельного алгоритма для модели общей циркуляции атмосферы. Сообщения по прикладной математике. М.: ВЦ РАН, 2004. 23 с.
3. Быков С.В., Пархоменко В.П. Моделирование циркуляции атмосферы, океана и эволюции морского льда в Арктическом регионе. Сообщения по прикладной математике. М.: ВЦ РАН, 2002. 36 с.
4. Белов П.Н., Борисенков Е.П., Панин Б.Д. Численные методы прогноза погоды. Л: Гидрометеиздат, 1989. 375 с.
5. Arakawa A., Lamb V., Computational Design of the Basic Dynamical Processes of the UCLA General Circulation Model // Methods in Comp. Phys. 1977, Vol.17, P. 173-265.
6. Parkhomenko V.P. Parallel Atmospheric General Circulation Model Code Analysis and Optimization // Research activities in atmospheric and oceanic modelling. World Meteorological Organization Geneva Switzerland, 2005, Vol. 35, P. 3.19 – 3.20.